

Content Liberation Update, August 2018

Rebekah Kati
Institutional Repository Librarian
University of North Carolina at Chapel Hill

Introduction

In 2018, the Open Access Implementation Team (the Team) conducted three Content Liberation pilot projects to ingest UNC-authored content into the Carolina Digital Repository (CDR). A summary of each project is available in [Content Liberation Overview](#).

The following report summarizes our current progress on these projects and suggests recommendations for progress in the next academic year. A follow-up report will be available in August 2019.

Author Citations

This pilot project has been largely on hold due to the purchase of 1science's 1foldr report. The report identified 47,281 green and gold open access articles that had been published by UNC researchers from January 1980-May 2018. In addition to article identification, the report also provided:

- Metadata for all articles, including DOIs and PMIDs when available
- SHERPA/RoMEO categories, summaries and link to policies for each article
- Link to download full text of articles from publisher's site, PubMed, author or institutional websites
- Macros to download all articles in the spreadsheet at once
- Macros to upload articles to an Eprints, DSpace or Digital Commons repository

The content identified by the 1foldr report was freely available and could have been obtained by our team. However, the process would not have been scalable for our small staff and would not have been feasible to accomplish within the same timeframe.

After the new CDR platform launch, the Team will load content from 2016 to the present. The remaining content will be loaded as staff time permits. Content which does not fall under the open access policy needs to be vetted closely by staff to determine the version that the publisher's policy allows us to deposit. 1foldr's data and automation will speed up this work significantly.

While 1foldr contributed a huge number of potential deposits to our backfile, the Author Citations project holds value to identify articles going forward. The Author Citations project identifies both paywalled and open access content. We are continuing to collect metadata for UNC-authored articles and will develop an outreach strategy after the new CDR platform launch.

CV Review

The Team developed and documented the workflow for locating publications, comparing the publication list with the SHERPA/RoMEO API to determine which publications were eligible for deposit in the CDR, and batch uploading deposits in the CDR.

One of the goals for the CV Review Project was to gauge the feasibility of this process for increasing the number of Humanities and Social Science faculty deposits. However, we did not receive enough data to make any conclusions about the amount of time each CV review takes to complete, nor how many publications we could expect on average from faculty across various disciplines.

Another goal was to determine if this was a useful service that UNC Libraries should continue to offer faculty, particularly as an outreach strategy to help faculty learn about Open Access options for their discipline. The Team was invited and welcomed to multiple departmental meetings in the Spring 2018 Semester and has already scheduled several more faculty visits for the upcoming Fall 2018 semester. Despite the labor-intensive process, we do feel this is a valuable service that the Libraries should continue to offer, and it will remain part of the outreach plan.

Highly Cited Authors

The Team has identified 56 UNC researchers from five years of Highly Cited Authors lists. We determined that these researchers had written a total of 842 highly cited articles that might be eligible for deposit in the CDR. The Team defined “highly cited article” as an article which received over 300 citations in Scopus.

As of mid-August 2018, the Team has vetted all 842 articles for eligibility, contacted 21 out of 56 researchers and deposited 286 articles. We received positive responses from the faculty and only one has declined to have his articles deposited. In one instance, an inquiry to a highly cited researcher led to an opportunity to publish more of his consortium’s work into the CDR.

The Team performed an assessment to determine whether the 842 articles were already available openly prior to deposit in the CDR. Using Unpaywall’s Simple Query Tool, we found that:

- 386 articles were not available openly in a repository or on the publisher’s website (45.84% liberated)
- 456 were already available openly in a repository or on the publisher’s website (54.16%)

Additionally, we also determined whether the 842 articles were available on an academic social network or author’s institutionally affiliated website:

- 396 articles were not available on an academic social network or an author’s institutionally affiliated website (47.03%)
- 446 were available on academic social networks or an author’s institutionally affiliated website (52.97%)

We will continue to contact the remaining researchers and deposit their work. Although much of the scholarship is open elsewhere, we feel that it is important enough to preserve in the CDR. Additionally,

we have raised awareness of the CDR to highly influential faculty, which could lead to increased deposits in the future.